EU policy reform simulation based on panel data estimation of on- and off-farm labour supply equations for Dutch dairy farmers

D.L. Ooms[1]

A.R. Hall[2]

[1]Agricultural Economics and Rural Policy Group

Department of Social Sciences

Wageningen University

Hollandseweg 1

6706 KN Wageningen

The Netherlands

e-mail: daan.ooms@wur.nl

[2]Department of Economics

North Carolina State University

Raleigh, NC USA

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July 24-27, 2005

Abstract

This research focuses on the estimation of labour supply equations for Dutch dairy farmers that are suitable for policy simulations. Data availability leads to the fact that we can not estimate structural labour supply equations. We show how to derive reduced form equations suitable for policy simulations. In this research we use the panel data sample selection estimation approach of Kyriazidou (1997) and Wooldridge (1995) to estimate the off-farm labour supply equation. The two lead to different estimation results and different simulation results based on these.

Keywords: Econometrics, Panel Data, Sample Selection, Labour Supply, CAP Reform.

JEL: C23, C24, C51, C53, D13, J22, Q12, Q18.

1. Introduction

On June 26 2003 the European ministers of agriculture agreed on a reform of the Common Agricultural Policy (CAP) (European Commission, 2003). For the Dutch dairy sector the reform entails a reduction in milk price, which is partly compensated by the introduction of direct income payments and a small increase in quota amount. Part of the policy reform is cross-compliance, this implies that (part of) direct payments can only be obtained if the farmer complies with certain criteria. Given the importance of the Dutch dairy sector it is important to assess the effects of the 2003 CAP reform. Direct income payments enter the farm household as non-labour income. The effects of non-labour income on time allocation Huffman and El-Osta (1998) and Woldehanna et al. (2000), amongst others, find, are contradicting. This indicates that the qualitative effect of direct income payments on time allocation can not be predicted beforehand and that estimation of this effect is needed.

In The Netherlands, the Dutch Agricultural Research Institute (LEI) creates an extensive farm level panel data set, containing production and consumption variables, of farm households. Panel data estimation methods control for unobserved heterogeneity between farms. Time-series and cross-section studies not controlling for this heterogeneity run the risk of obtaining biased results (Baltagi, 2001). For this reason we will use panel data methods. Unfortunately the data set does not contain individual off-farm hourly wage. For this reason we can not estimate structural on-and off-farm labour supply equations. We will show that we can estimate reduced form labour supply equations that are suitable for 2003 CAP reform simulations.

Large part of Dutch dairy farm households does not supply off-farm labour. Estimating an off-farm labour supply equation based on only the farm households that do supply off-farm labour can lead to sample selection and has to be taken into account in estimation (Heckman, 1979). Both Kyriazidou (1997) and Wooldridge (1995), amongst others, introduce panel data sample selection model estimation approaches. To the best of our knowledge these sample selection approaches have not been used in agricultural labour economics.

The purpose of this paper is to estimate reduced form on-farm and off-farm labour supply equation for Dutch dairy farmers using panel data estimation techniques and taking possible sample selection in the off-farm labour supply equation into account. These reduced form equations are used to determine the effects of the EU 2003 dairy policy reforms.

The remainder of this paper contains the following. In section 2 a theoretical derivation of on- and off-farm labour supply equations is given. In section 3 we derive reduced form labour supply equations suitable for policy simulations. Section 4 describes the data used for estimation and simulation. Section 5 gives descriptions of two linear panel data estimation approaches used and Section 6 gives descriptions of two panel data sample selection estimation approaches used. Section 7 describes the estimation results. Section 8 describes the policy simulations of which the results are described in Section 9. Finally, section 10 gives a brief summary and conclusions.

2. Theoretical model

The following theoretical model is based on the household utility model of Huffman (1980). Labour supply decisions of dairy farm household i at time t are assumed to be the result of maximising utility (u_{it}) received from consuming goods and services (c_{it}) and home time ($h_{h,it}$) given a vector of utility shifting household characteristics

($\mathbf{z}_{u,it}$) and a vector of other variables influencing the households' decision making environment ($\mathbf{o}_{it}$),

$$u_{it} = u(c_{it}, h_{h,it}; \mathbf{z}_{\mathbf{u},it}, \mathbf{o}_{it}) \qquad (1)$$

where $u(\ )$ is a utility function that is the same for all households. Differences between the utility levels of households come from the different choices made with respect to the elements of the utility function. Total time endowment (h_{it}^0) is allocated between farm labour ($h_{f,it}$), off-farm labour ($h_{of,it}$) and home time. Which results in the time constraint:

$$h_{it}^0 = h_{f,it} + h_{of,it} + h_{h,it}, \quad h_{of,it} \geq 0. \qquad (2)$$

The time constraint is a strict equality because home time is defined to be the difference between total time and labour time. Home time consists of leisure, household work, etc. A non-negativity constraint is imposed on off-farm labour because it may be zero. Throughout, we assume all prices to be the same for all households and only differ between time periods. Dairy farmers in The Netherlands produce milk ($q_{m,it}$) and one or more other output ($q_{o,it}$). For this production the farmer uses variable input (g_{it}), cattle (m_{it}), farm labour and factor inputs ($\mathbf{z}_{\mathbf{q},it}$). Since milk output is produced under a quota system it is assumed fixed on the short term. We assume farm households minimise short-term costs given prices of inputs (v_t), the price of cattle ($v_{m,t}$), prices of other outputs ($p_{o,t}$), farm labour, factor inputs and milk output. Other outputs generate revenue and are therefore seen as negative costs in the following short-term cost function $k(\ )$:

$$k(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}) = \\ \min_{g_{it}, m_{it}, q_{o,it}} \{ g_{it} v_t' + m_{it} v_{m,t} - q_{o,it} p_{o,t}' \mid h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it} \} \qquad (3)$$

This cost function is assumed to be continuous and twice differentiable non-decreasing and concave in input prices and the price of cattle, non-increasing and convex in other output prices and linear homogeneous in all prices. The shadow price of farm labour $s_{f,it}(\;)$ is:

$$\frac{\partial k\left(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}\right)}{\partial h_{f,it}} = s_{f,it}\left(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}\right) \quad (4)$$

The shadow price of farm labour is the marginal cost of using an extra unit of labour in production. Since farm labour is owned by the farm this marginal cost for production is equal to the marginal revenue of labour for the farm household. The shadow price of labour is the price at which the internal market of farm labour supply clears. The equations for the shadow prices of factor inputs and milk output are similar to equation (4). Because we are mainly interested in labour in this paper we omit the explicit equations for these other shadow prices. Farm income ($y_{f,it}$) equals milk revenue minus costs:

$$y_{f,it} = p_{m,t} q_{m,t} - k\left(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}\right) \quad (5)$$

Off-farm labour income is defined as off-farm labour time times off-farm wage ($w_{of,it}$):

$$y_{of,it} = w_{of,it} h_{of,it} \quad (6)$$

Notice that wages are farm and time specific. The value of household consumption is defined as the product of consumption goods and services with the price of consumption goods and services ($p_{c,t}$). Consumption is constrained by total income. Total income consists of farm income, off-farm labour income and other income ($y_{o,it}$):

$$y_{f,it} + y_{of,it} + y_{o,it} = c_{it} p_{c,t}' \quad (7)$$

Other income contains, amongst others, subsidies in the form of income transfers, whereas the EU milk price support is part of farm income through the price variables in equation (5). Combining equations (5) to (7) results in:

$$p_{m,t}q_{m,t} - k\left(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}\right) + w_{of,it}h_{of,it} + y_{o,it} = c_{it}\, p_{c,t}' \qquad (8)$$

We assume the household maximises (1) subject to (2) and (8) by choosing the elements of the choice set $c_{it}, h_{h,it}, h_{f,it}, h_{of,it}, g_{it}, q_{o,it}$. The Kuhn-Tucker first-order conditions are:

$$\frac{\partial u}{\partial c_{it}} = \lambda_1 p_{c,t}, \qquad (9)$$

$$\frac{\partial u}{\partial h_{h,it}} = \lambda_2, \qquad (10)$$

$$\lambda_1 \frac{\partial k}{\partial h_{f,it}} - \lambda_2 = 0, \qquad (11)$$

$$\lambda_1 w_{of,it} - \lambda_2 \le 0, \quad h_{of,it} \ge 0, \quad h_{of,it}\left(\lambda_1 w_{of,it} - \lambda_2\right) = 0, \qquad (12)$$

$$\lambda_1 \frac{\partial k}{\partial g_{it}} = 0, \qquad (13)$$

$$\lambda_1 \frac{\partial k}{\partial q_{o,it}} = 0 \qquad (14)$$

plus equations (2) and (8) where λ_1 is the marginal utility of income and λ_2 is the marginal utility of time. If an interior solution exists (i.e. off-farm labour supply is non-zero) the first part of equation (12) holds as equality. In this case the first order conditions can be solved to yield:

$$\frac{\partial u\left(c_{it}, h_{h,it}; \mathbf{z}_{\mathbf{u},it}, \mathbf{o}_{it}\right)/\partial c_{it}}{\partial u\left(c_{it}, h_{h,it}; \mathbf{z}_{\mathbf{u},it}, \mathbf{o}_{it}\right)/\partial h_{h,it}} = -\frac{p_{c,t}}{w_{of,it}} \qquad (15)$$

and

$$\frac{\partial k\left(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}\right)}{\partial h_{f,it}} = w_{of,it}\,. \tag{16}$$

If a farm household only works on the farm, the first part of equation (12) does not hold as equality and:

$$\frac{\partial k\left(v_t, v_{m,t}, p_{o,t}, h_{f,it}, \mathbf{z}_{\mathbf{q},it}, q_{m,it}\right)}{\partial h_{f,it}} > w_{of,it} \tag{17}$$

Equation (15) implies that the marginal rate of substitution between leisure and consumption goods is equal to the ratio of the consumption good prices to the wage rate. Equation (16) implies that, if off-farm labour is supplied, the marginal product of farm labour is equal to off-farm wage. Equation (17) implies that, if no off-farm labour is supplied, the marginal product of farm labour is strictly larger than off-farm wage. The left-hand side of equation (16) and (17) is the shadow price of labour used on the farm. This shadow price, the marginal product of labour, does not depend on the output price of milk (see Appendix A for a graphical amplification). From equation (2) and (8) till (14) we derive the reduced form functions for off-farm labour supply $l_{of}(\ )$ and on-farm labour supply $l_f(\ )$. These are functions of all variables in equation (1) to (8) except the variables in the choice set. This results in:

$$h_{of,it} = l_{of}\left(w^n_{of,it}, p^n_{c,t}, p^n_{m,t}, v^n_{m,t}, q_{m,it}, p^n_{o,t}, y^n_{o,it}, \mathbf{z}_{\mathbf{u},it}, \mathbf{z}_{\mathbf{q},it}, \mathbf{o}_{it}, t^0_{it}\right), \tag{18}$$

$$h_{f,it} = l_f\left(w^n_{of,it}, p^n_{c,t}, p^n_{m,t}, v^n_{m,t}, q_{m,it}, p^n_{o,t}, y^n_{o,it}, \mathbf{z}_{\mathbf{u},it}, \mathbf{z}_{\mathbf{q},it}, \mathbf{o}_{it}, t^0_{it}\right). \tag{19}$$

Here, the superscript *n* indicates that the corresponding variable is normalised by the price of variable inputs. This is done to impose linear homogeneity in prices and income. Equations (18) and (19) contain milk price, milk output (that is equal to milk quota in The Netherlands) and other income. These are all variables that are influenced by the 2003 CAP reform for dairy farmers. For this reason equation (18) and (19) are suitable for our goal of constructing models for 2003 CAP policy simulation.

3. Empirical model

Equation (18) and (19) show that labour supply is expressed in time units and is, amongst others, explained by individual wages. The data set of Dutch dairy farmers at our disposal contains on-farm hours. However, it does not contain off-farm hours and individual wages. This section explains how we deal with these limitations and results in empirical specifications for equation (18) and (19).

3.1 Off-farm labour supply data

Instead of off-farm hours, our data set contains off-farm income. This paragraph explains how it is possible to estimate the parameters in the off-farm labour supply equation, using off-farm income instead of hours. To this end we choose the functional form in (18) as:

$$\ln[h_{of,it}] = \beta_i + \beta_w \ln[w_{of,it}] + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\beta}_{-\mathbf{w}} + \varepsilon_{it}\,. \tag{20}$$

Here, $w_{of,it}$ represents the wage and $\mathbf{x}_{-\mathbf{w},it}$ is the row vector of all explanatory variables except off-farm wage. ε_{it} is an error term with expectation zero. β_i is a farm specific effect, β_w is the parameter associated with the log of wage and the column vector $\boldsymbol{\beta}_{-\mathbf{w}}$ contains parameters associated with the other explanatory variables. For notational convenience, the content of $\mathbf{x}_{-\mathbf{w},it}$ is not specified the equations in this section.

Off-farm income $y_{of,it}$ is by definition:

$$y_{of,it} \equiv h_{of,it} w_{of,it}\,. \tag{21}$$

Taking the natural logarithm of (21) gives:

$$\ln y_{of,it} \equiv \ln h_{of,it} + \ln w_{of,it}\,. \tag{22}$$

Replacing $\ln h_{of,it}$ in (20) by equation (22) results in:

$$\ln y_{ofit} - \ln w_{of,it} = \beta_i + \beta_w \ln w_{of,it} + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\beta}_{-\mathbf{w}} + \varepsilon_{it} \tag{23}$$

and

$$\ln y_{ofit} = \beta_i + \breve{\beta}_w \ln w_{of,it} + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\beta}_{-\mathbf{w}} + \varepsilon_{it} \,. \tag{24}$$

Here, $\breve{\beta}_w = \beta_w + 1$. This derivation shows the possibility to estimate the parameters of equation (22), in which labour time is the dependent variable, using the logarithm of labour income as dependent variable instead. The only thing that has to be taken into account is that the estimated parameter on $\ln w_{of,it}$ is not β_w but $(\beta_w + 1)$.

3.2 Off-farm wage

Instead of farm specific off-farm wages, our data set contains national wages for labourers in the agricultural sector (w_t). Data on these wages differ between periods, not between farms. Using this national wage rate for individual farmers implies a measurement error (ζ_{it}). We assume:

$$w_{of,it} = w_t \zeta_{it} \quad \text{with } \zeta_{it} > 0 \tag{25}$$

Inserting (25) in (24) gives:

$$\ln[y_{of,it}] = \beta_i + \breve{\beta}_w \ln[w_t] + \breve{\beta}_w \ln[\zeta_{it}] + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\beta}_{-\mathbf{w}} + \varepsilon_{it} \tag{26}$$

We assume that the relation between the measurement error and the other explanatory variables is:

$$\ln[\zeta_{it}] = \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\alpha} + a_{it} \tag{27}$$

,where $E[a_{it} \mid \mathbf{x}_{-\mathbf{w},it}] = 0$. For variables in $\mathbf{x}_{-\mathbf{w},it}$ that are correlated with off-farm wage the corresponding parameter in $\boldsymbol{\alpha}$ is not zero. The interpretation of the parameters for these variables is altered by the existence of the measurement error. The same holds for the parameter associated with off-farm wage. For variables in $\mathbf{x}_{-\mathbf{w},it}$ that are not correlated with off-farm wage the corresponding parameter in $\boldsymbol{\alpha}$ is zero. These parameter estimates are not influenced by the measurement error under the condition that the variables that do explain off-farm wage are present in $\mathbf{x}_{-\mathbf{w},it}$. To control for the latter we insert education level and age in $\mathbf{x}_{-\mathbf{w},it}$. The equation to be estimated is:

$$\ln y_{ofit} = \beta_i + \breve{\beta}_w \ln[w_t] + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\beta}_{-\mathbf{w}} + \tau_{it} \tag{28}$$

Here, $\tau_{it} \equiv \breve{\beta}_w \ln \zeta_{it} + \varepsilon_{it}$. $\mathbf{x}_{-\mathbf{w},it}$ contains milk price; quota amount and external income. These variables are influenced by the 2003 CAP reform. From the reasoning above follows that the parameter interpretation for off-farm wage; education level and age is altered. However the parameter interpretation for the policy variables is not. Therefore, estimating equation (28) does not lead to a correct off-farm labour supply equation, but the estimated equation is suitable for our policy simulations. Equation (28) in which $\mathbf{x}_{-\mathbf{w},it}$ contains the explanatory variables in (18) except off-farm wage is used in estimation.

From (19) follows that off-farm wage is also an explanatory variable in the on-farm labour supply function. Therefore, in this function we also have to deal with the fact that we do not have data on farm specific off-farm wages. To this end we choose the functional form in (19) comparable to the functional form in (18) given in equation (20):

$$\ln[h_{f,it}] = \gamma_i + \gamma_w \ln[w_{of,it}] + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\gamma}_{-\mathbf{w}} + \nu_{it} \tag{29}$$

Here, ν_{it} is an error term with expectation zero. γ_i is a farm specific effect, γ_w is the parameter associated with the log of wage and the column vector $\boldsymbol{\gamma}_{-\mathbf{w}}$ contains parameters associated with the other explanatory variables. Inserting (26) in (29) gives:

$$\ln[h_{fit}] = \gamma_i + \gamma_w \ln[w_t] + \mathbf{x}_{-\mathbf{w},it}\boldsymbol{\gamma}_{-\mathbf{w}} + \omega_{it} \tag{30}$$

Here, $\omega_{it} \equiv \gamma_w \ln \zeta_{it} + \nu_{it}$. Similar to the off-farm labour supply equation the parameter interpretation for off-farm wage; education level and age is altered. However the parameter interpretation for the policy variables is not. Therefore, estimating equation (30) does not lead to a correct on-farm labour supply equation, but the estimated equation is suitable for our policy simulations. Equation (30) in which $\mathbf{x}_{-\mathbf{w},it}$ contains the explanatory variables in (19) except off-farm wage is used in estimation.

4. Data

This section gives a description of the data used in estimation. The farm specific data come from the Dutch Agricultural Research Institute (LEI) unbalanced rotating panel data set of Dutch farms. A farm is classified to be a dairy farm if its returns consist for 50% or more of milk revenues. The data set consists of 6338 observations on 1307 farms. The period investigated is from 1987/88 until 1999/00. National data come from Statistics Netherlands (CBS). Off-farm labour is represented by off-farm income. The total number of family hours worked on the farm represents on-farm labour. Off-farm wage is represented by the national index of wages for agricultural hired labour. 1991 is the base year for this and subsequent indices. Price variables, influencing short-term farm income, are the milk price index, a Thornqvist price index for variable input, the price index for cattle, and a Thornqvist price index for other output. Variable input contains, amongst others, feed and veterinary costs. Cattle consists of cows aged one and older. Other output contains marketable crops, veal, pigs, poultry and other farm revenues. Quota is the amount of milk output a farmer is allowed to produce and is expressed in metric tonne. Other income is a monetary value. It includes, amongst other, income from externally allocated capital, income from social allowances and subsidies in the form of income transfers. These subsidies are mainly the acreage premium for maize. Most farmers produce more maize than they get subsidy for. From this follows that the premium does not influence maize production and the subsidy can be seen as an income transfer. Land is expressed in the number of hectares used by the farmer. Machinery is the average value of machinery over the year divided by the Thornqvist price index for machinery. Debt is the total value of short and long-term debt. Assets are represented by their value calculated by the LEI. Unemployment is expressed as the national unemployment rate. From the LEI data set it is possible to derive the importance of different activities as a percentage of total activity. This is based on output and the allocation of resources. The percentage for milk production is used as the specialisation rate of dairy farmers. Household variables used are number of household members; a dummy for the presence of a successor and a dummy variable indicating the education level of the head of the household. In table B.1 in Appendix B an overview of the dimension mean and standard deviation of the variables used is given.

5. Estimation of linear panel data models

Since we have panel data at our disposal we can use the extra information contained in panel data compared to cross-section data by using a linear panel data model estimation approach to estimate equation (30) in which $\mathbf{x}_{-\mathbf{w},it}$ contains the explanatory variables in (19) except off-farm wage. A linear panel data model is given by:

$$h_{it} = \mathbf{x}_{\mathbf{1},it}\boldsymbol{\theta}_{\mathbf{1}} + \mathbf{x}_{\mathbf{2},i}\boldsymbol{\theta}_{\mathbf{2}} + a_i + v_{it} \tag{31}$$

where h_{it} is the dependent variable, $\mathbf{x}_{\mathbf{1},it}$ is a vector of observable explanatory variables that vary both over farms and time with corresponding vector of unobservable parameters $\boldsymbol{\theta}_{\mathbf{1}}$, $\mathbf{x}_{\mathbf{2},it}$ is a vector of observable explanatory variables that vary over farms but are constant over time $\boldsymbol{\theta}_{\mathbf{2}}$, a_i is an unobservable farm effect and v_{it} is an unobservable error term. Two linear panel data model estimation methods are used in this research: the fixed effects method and the Mundlak's (1978) method.

5.1 The fixed effects method

The fixed effects method (FE) transforms the elements of (31) to eliminate the individual specific effects. The transformation performed subtracts the average of equation (31) over time from equation (31) to result in:

$$\ddot{h}_{it} = \ddot{\mathbf{x}}_{\mathbf{1},it}\boldsymbol{\theta}_{\mathbf{1}} + \ddot{v}_{it} \tag{32}$$

Here, $\ddot{h}_{it} \equiv h_{it} - \bar{h}_{it}$ and $\bar{h}_i = \frac{1}{T_i}\sum_{t=1}^{T} h_{it}$. T_i is the number of periods individual *i* is in the data set. T is the total number of periods for which data are available. If individual *i* is not in the data set at time *t*, $h_{it} = 0$. The other variables are transformed in a similar way. This transformation is called the within transformation. The within transformation for the individual specific effect is zero by definition. The same holds for the time invariant variables. Fe results in a consistent estimate of $\boldsymbol{\theta}_1$, however it does not produce an estimate for $\boldsymbol{\theta}_2$. See e.g. Wooldridge (2002) for a more elaborate explanation of FE.

5.2 Mundlak (1978)

Mundlak deals with the individual specific effects in (31) by replacing them by the average values over time of the explanatory variables resulting in:

$$h_{it} = \mathbf{x}_{1,it}\boldsymbol{\theta}_1 + \mathbf{x}_{2,i}\boldsymbol{\theta}_2 + \bar{\mathbf{x}}_{1,i}\boldsymbol{\rho} + v_{it} \tag{33}$$

Here $\boldsymbol{\rho}$ is a parameter vector to be estimated. The parameter vectors are estimated by pooled regression of equation (33). The Munlak (1978) approach does not transform variables. Therefore, with this approach it is possible to obtain parameter estimates for the time invariant variables.

6. Estimation of panel data models with sample selection

Not all farmers in The Netherlands supply off-farm labour. It is assumed that the group of farmers that do supply off-farm labour is not a representative sample of all farmers. This calls for a sample selection estimation approach. Since we have panel data at our disposal we can use the extra information contained in panel data compared to cross-section data by using a panel data sample selection model estimation approach to estimate equation (28) in which $\mathbf{x}_{-\mathbf{w},it}$ contains the explanatory variables in (18) except off-farm wage. This section describes the estimation methods proposed by Kyriadzidou (1997) and Wooldridge (1995), both of which are employed in the empirical analysis.

Off-farm labour is only observed for those households that choose to supply it. We therefore introduce an indicator variable d_{it} that takes the value one if household i supplies off-farm labour in period t. It is assumed that this decision depends on a vector of explanatory variables $\mathbf{m}_{it}$ via

$$d_{it} = I\{\mathbf{m}_{it}\boldsymbol{\gamma} + b_i - u_{it} \geq 0\} \tag{34}$$

where $I\{\cdot\}$ is an indicator function that takes the value one if the event in the curly brackets occurs but is zero otherwise, γ is a vector of unknown parameters, b_i is an unobservable farm effect, and u_{it} is an unobservable error term. If $d_{it} = 1$ and so off-farm labour is supplied then the log of off-farm labour supply, y_{it}, is assumed to be generated by

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,i}\boldsymbol{\beta}_2 + c_i + v_{it} \tag{35}$$

where $\mathbf{x}_{1,it}$ is a vector of observable explanatory variables that vary both over farms and time with corresponding vector of unobservable parameters $\boldsymbol{\beta}_1$, $\mathbf{x}_{2,it}$ is a vector of observable explanatory variables that vary over farms but are constant over time with corresponding vector of unobservable parameters $\boldsymbol{\beta}_2$, c_i is an unobservable farm effect and v_{it} is an unobservable error term.

Two problems arise in the estimation of equation (35). First, the individual effect, c_i, is unobserved. Second, there is a potential sample selection bias if the selection equation, equation (34), does not select a random sample from the underlying population. As we now describe, Kyriazidou (1997) and Wooldridge (1995) take different approaches to address these problems in their estimation methods.

6.1 Kyriazidou's (1997) method:

Kyriazidou (1997) circumvents the presence of the individual effect by basing the estimation on a differenced version of (35). The errors are then assumed to satisfy a conditional exchangeability property so that the potential sample selection bias in the differenced version of the off-farm labour supply equation can be circumvented via a weighted least squares estimation. An advantage of this approach is that it does not require the specification of a parametric form for the error distribution. The disadvantage is that it only provides estimates of $\boldsymbol{\beta}_1$, that is the parameters on the time varying explanatory variables in the off-farm labour supply equation.

To present Kyriazidou's method, it is useful to define a vector that contains all observed and unobserved variables associated with farm *i* in periods *t* and *s*, that is $\mathbf{w}_{its} = (\mathbf{m}_{it}, \mathbf{m}_{is}, \mathbf{x}_{1,it}, \mathbf{x}_{1,is}, \mathbf{x}_{2,i}, b_i, c_i)$. Without loss of generality, it is assumed that $t > s$. Kyriazidou's method exploits pairs of observations (i,t) and (i,s) for which $d_{it} = d_{is} = 1$ that is observations on a farm household that supplies off-farm labour in two different time periods. Kyriazidou (1997) shows that (35) can be re-written as

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,i}\boldsymbol{\beta}_2 + c_i + \lambda_{its} + \tilde{v}_{it} \tag{36}$$

where

$$\lambda_{its} = E[v_{it} \mid d_{it} = 1, d_{is} = 1, w_{its}] \tag{37}$$

and $\tilde{v}_{it} = v_{it} - \lambda_{its}$ has zero expectation by construction. The variable λ_{it} is non-zero if the selection equation does not yield a random sample from the underlying population.

From (36), it follows that

$$y_{it} - y_{is} = (\mathbf{x}_{1,it} - \mathbf{x}_{1,is})\boldsymbol{\beta}_1 + \lambda_{its} - \lambda_{ist} + \tilde{z}_{its} \tag{38}$$

where $\tilde{z}_{its} = \tilde{v}_{it} - \tilde{v}_{is}$ has zero expectation conditional on $\mathbf{w}_{its}$ by construction. Kyriazidou observes that if both $(v_{it}, v_{is}, u_{it}, u_{is})$ and $(v_{is}, v_{it}, u_{is}, u_{it})$ are identically distributed conditional on $\mathbf{w}_{its}$ [1] and also $\mathbf{x}_{it}\boldsymbol{\gamma} = \mathbf{x}_{is}\boldsymbol{\gamma}$ then $\lambda_{its} = \lambda_{ist}$. This means that, for these individuals, equation (38) reduces to

$$y_{it} - y_{is} = (\mathbf{x}_{\mathbf{1},it} - \mathbf{x}_{\mathbf{1},is})\boldsymbol{\beta}_{\mathbf{1}} + \tilde{z}_{its} \tag{39}$$

which is free of both individual effects and the terms due to sample selection. A consistent estimator of $\boldsymbol{\beta}_{\mathbf{1}}$ could, in principle, be obtained via Ordinary Least Squares (OLS) estimation of (39) using those observations for which $d_{it} = d_{is} = 1$ and $\mathbf{x}_{it}\boldsymbol{\gamma} = \mathbf{x}_{is}\boldsymbol{\gamma}$.

While the conditional exchangeability assumption is plausible in many circumstances and it is likely that there are farms which provide off-farm labour in more than one period, it is unlikely that $\mathbf{x}_{it}\boldsymbol{\gamma} = \mathbf{x}_{is}\boldsymbol{\gamma}$ holds for those concerned as $\mathbf{x}_{it}$ typically contains time varying variables. Therefore, the aforementioned OLS estimation is infeasible. Instead, Kyriazidou (1997) proposes to estimate (39) based on observations for which $d_{it} = 1$ and $d_{is} = 1$ using Weighted Least Squares (WLS) with weights inversely proportional to $|\mathbf{m}_{it}\boldsymbol{\gamma} - \mathbf{m}_{is}\boldsymbol{\gamma}|$. Specifically, the WLS estimator is

$$\hat{\boldsymbol{\beta}}_{\mathbf{1}}(k) = \left[\sum_{i=1}^{n}\left\{\frac{1}{T_i - 1}\right\}\sum_{s<t}\hat{\psi}_{its,n}(\mathbf{x}_{\mathbf{1},it} - \mathbf{x}_{\mathbf{1},is})(\mathbf{x}_{\mathbf{1},it} - \mathbf{x}_{\mathbf{1},is})'\right]^{-1} \times \left[\sum_{i=1}^{n}\left\{\frac{1}{T_i - 1}\right\}\sum_{s<t}\hat{\psi}_{its,n}(\mathbf{x}_{\mathbf{1},it} - \mathbf{x}_{\mathbf{1},is})(y_{it} - y_{is})'\right] \tag{40}$$

where n is the number of farm households, T_i is the number of periods that farm i is in the sample, $\sum_{s<t}$ denotes the sum over all combinations of s and t for which $s,t = 1,2,...,T_i$ and $s<t$, and $\hat{\psi}_{its,n}$ are the weights. Kyriazidou proposes using weights of the form

[1] This assumption is termed "conditional exchangeability" by Kyriazidou (1997). Note that the event $d_{it} = d_{is} = 1$ is equivalent to $\mathbf{x}_{it}\boldsymbol{\gamma} + b_i \geq u_{it}, \mathbf{x}_{is}\boldsymbol{\gamma} + b_i \geq u_{is}$ and hence the probability of its occurrence is governed by u_{it} and u_{is}.

$$\hat{\psi}_{its,n} = \frac{1}{h_n} K\left(\frac{(\mathbf{m}_{it} - \mathbf{m}_{is})\hat{\boldsymbol{\gamma}}_n}{h_n} \right) \qquad (41)$$

where $K(.)$ is a kernel function and $\hat{\boldsymbol{\gamma}}_n$ is a consistent estimator of $\boldsymbol{\gamma}$.[2] In the estimations reported below, we set $K(.)$ equal to the probability density function of the standard normal distribution[3], $h_n = n^{-1/5}$, which is the bandwidth value corresponding to the standard normal distribution that maximises the rate of convergence in distribution of $\hat{\boldsymbol{\beta}}_{\mathbf{1}}(k)$ and $\hat{\boldsymbol{\gamma}}_n$ is the fixed effects logit estimator of $\boldsymbol{\gamma}$.[4] Kyriazidou (1997) proves that $\hat{\boldsymbol{\beta}}_{\mathbf{1}}(k)$ is asymptotically biased but that this bias can be removed by transforming the estimator as follows,

$$\hat{\boldsymbol{\beta}}_{\mathbf{1},c}(k) = \frac{\hat{\boldsymbol{\beta}}_{\mathbf{1}}(k) - n^{(1-\delta)2/5}\hat{\boldsymbol{\beta}}_{\mathbf{1}}(k;\delta)}{1 - n^{(1-\delta)2/5}} \qquad (42)$$

where $\hat{\boldsymbol{\beta}}_{\mathbf{1}}(k;\delta)$ is the estimator of $\boldsymbol{\beta}_{\mathbf{1}}$ in (40) evaluated with $h_n = n^{-\delta/5}$ and any δ satisfying $0 < \delta < 1$. In our calculations, we set $\delta = 0.1$. Kyriadzidou(1997) shows that $\hat{\boldsymbol{\beta}}_{\mathbf{1},c}(k)$ is asymptotically unbiased and normally distributed.

6.2 Wooldridge's (1995) method:

Rather than eliminate the individual effect via a transformation, Wooldridge (1995) models c_i as an explicit function of the explanatory variables in the fashion proposed by Mundlak (1978) and Chamberlain (1982). The potential sample selection bias is circumvented by including the Mill's ratio as an additional regressor in the off-farm labour supply equation in the spirit of Heckman (1979). For the latter device to be successful, the errors to the off-farm labour supply and selection equations must be jointly normally distributed. Therefore, in contrast to Kyriazidou's method, Wooldridge's approach requires explicit parametric assumptions about the individual effect and the error distribution. The relative advantage is that if these assumptions are

[2] See Bierens (1987) for an overview of different kernel functions and their properties.
[3] See *inter alia* Greene (1997)[p.68].
[4] See Wooldridge (2002)[p.491]. Note that to evaluate $(\mathbf{m}_{it} - \mathbf{m}_{is})\hat{\gamma}_n$ we only require a consistent estimator of the parameters on time varying explanatory variables in the selection equation.

correct then Wooldridge's method yields consistent estimators of $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$, that is the parameters on both the time varying and non-time varying explanatory variables.

In both equations, the individual effect is replaced by a linear combination of the means of the time varying explanatory variables. To make this substitution in the selection equation, it is necessary to define first a partition of $\mathbf{m}_{it}$ into time varying and non-time varying variables, that is $\mathbf{m}_{it} = (\mathbf{m}_{1,it}, \mathbf{m}_{2,it})$. The selection equation becomes

$$d_{it} = I\{\mathbf{m}_{1,it}\boldsymbol{\gamma} + \overline{\mathbf{m}}_{1,i}\boldsymbol{\pi} - u_{it} \geq 0\} \tag{43}$$

where $\overline{\mathbf{m}}_{1,i} = T_i^{-1}\sum_{t=1}^{T_i}\mathbf{m}_{1,it}$ and $\boldsymbol{\pi}$ is a vector of unknown parameters. Assuming that off-farm labour is supplied $(d_{it} = 1)$, the off-farm labour supply equation can be written as

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,i}\boldsymbol{\beta}_2 + \overline{\mathbf{x}}_{1,i}\boldsymbol{\kappa} + \eta\lambda_{it}(\boldsymbol{\gamma}, \boldsymbol{\pi}) + e_{it} \tag{44}$$

where $\overline{\mathbf{x}}_{1,i} = T_i^{-1}\sum_{t=1}^{T_i}\mathbf{x}_{1,it}$, $\boldsymbol{\kappa}$ is a vector of unknown parameters, and

$$\lambda_{it}(\boldsymbol{\gamma}, \boldsymbol{\pi}) = \frac{\phi(\mathbf{x}_{it}\boldsymbol{\gamma} + \overline{\mathbf{x}}_i\boldsymbol{\pi})}{\Phi(\mathbf{x}_{it}\boldsymbol{\gamma} + \overline{\mathbf{x}}_i\boldsymbol{\pi})} \tag{45}$$

Under the assumption that u_{it}, v_{it} are jointly normally distributed conditional on $\{\mathbf{m}_{it}, \overline{\mathbf{m}}_{1,i}, \mathbf{x}_{1,it}, \mathbf{x}_{2,i}, \overline{\mathbf{x}}_{1,i}\}$, the error term e_{it} satisfies $E[e_{it} \mid \mathbf{x}_{it}, \overline{\mathbf{x}}_i, \mathbf{m}_{it}, \overline{\mathbf{m}}_{1,i}] = 0$. Therefore, if $\boldsymbol{\gamma}$ and $\boldsymbol{\pi}$ were known (and so $\lambda_{it}(\boldsymbol{\gamma}, \boldsymbol{\pi})$ were calculable) then OLS estimation of (44) based on those observations for which $d_{it} = 1$ would yield a consistent estimator of $(\boldsymbol{\beta}', \boldsymbol{\kappa}', \eta)'$. In general, $\boldsymbol{\gamma}$ and $\boldsymbol{\pi}$ are unknown, and so the latter estimation is infeasible. To circumvent this problem, Wooldridge proposes obtaining preliminary estimates of the selection equation parameters, $(\hat{\boldsymbol{\gamma}}', \hat{\boldsymbol{\pi}}')'$ say, from a pooled probit estimation of (43), and then using these estimates to obtain the sample analogue to the Mill's ratio. Estimates of $(\boldsymbol{\beta}', \boldsymbol{\kappa}', \eta)'$ are then obtained via OLS regression of y_{it}

on $\mathbf{x}_{1,it}, \mathbf{x}_{2,i}, \bar{\mathbf{x}}_{1,i}$ and $\lambda_{it}(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\pi}})$ based on the sample of observations for which off-farm labour is supplied, that is $\{(i,t); d_{it} = 1\}$. Wooldridge shows that these estimates are consistent and asymptotically normal.

7. Estimation results

This section gives the estimation results for the off-farm labour supply equation (28) estimated with both the Kyriazidou (1997) and Wooldridge (1995) approach. Also the estimation results for the on-farm labour supply equation (31) estimated with FE and the Mundlak (1978) approach are given. The results for the model consisting of both the on- and off-farm labour supply equations is presented and described in such a way that the estimation methods used are based on comparable treatments of the individual specific effects. This results in a model estimated with Kyriazidou (1997) and FE and a model estimated with Wooldridge (1995) and Mundlak (1978). Remember that both the Kyriazidou (1997) and the Wooldridge (1995) estimation approaches require estimation of a binary choice model for off-farm labour supply. These results are not of primary interest in this paper. Therefore, we describe them shortly in Appendix C.

Sample selection estimation approaches require that we have at least one explanatory variable in the off-farm labour decision equation that does not appear in the off-farm labour supply equation. We choose to use on-farm specialisation in milk production as the variable that does appear in the on-farm labour supply and the off-farm labour decision equation, but not in the off-farm labour supply equation. The results of Weiss and Briglauer (2000) suggest that off-farm labour is a diversification choice in the reduction of risk. A high on-farm specialisation might increase the propensity to work off-farm to reduce risk. We assume that this effect works mainly through the diversification decision and less through the amount of diversification. Based on this assumption we only include on-farm specialisation in the off-farm labour decision equation and not in the off-farm labour supply equation. An effect of specialisation on on-farm labour use is an often-found result in economic research.

Table 1 gives the estimation results for the off- and on-farm labour supply equations for both the Kyriazidou (1997)/FE and the Wooldridge (1995)/Mundlak (1978) estimation approach.

Table 1: Estimation results for labour supply equations.

Estimation approach:	Kyriazidou (1997)		FE		Wooldridge (1995)		Mundlak (1978)	
Dependent variable:	Log off-farm labour		Log on-farm labour		Log off-farm labour		Log on-farm labour	
	Est	t-stat	est	t-stat	Est	t-stat	Est	t-stat
Constant					12.65	2.16*	5.91	13.92*
Log off-farm wage	2.80	8.44*	-0.37	-13.36*	3.60	7.35*	-0.27	-6.18*
Milk price	-2.20	-3.31*	0.22	3.63*	1.22	0.74	0.14	1.37
Other output price	0.85	2.91*	-0.04	-1.35	-0.03	-0.05	-0.02	-0.55
Cattle price	-1.16	-1.94*	-0.01	-0.25	3.88	2.06*	-0.04	-0.50
Other income	0.13	0.08	-0.88	-7.03*	-3.97	-1.71*	-1.04	-4.97*
Quota	-0.89	-2.35*	0.21	6.91*	-3.04	-3.94*	0.22	1.91*
Land	-0.13	-0.27	0.27	6.94*	2.93	3.32*	0.14	2.02*
Buildings	-1.54	-5.02*	0.00	0.06	-2.18	-4.12*	-0.03	-1.12
Machinery	0.48	0.86	0.03	0.56	-0.83	-0.97	0.00	0.01
Debt over asset ratio	0.06	0.27	-0.08	-3.61*	0.70	1.57	0.03	0.99
Unemployment rate	-0.02	-1.16	0.00	0.91	0.04	1.37	0.00	0.31
Age	0.33	1.03	0.31	11.37*	-0.19	-3.06*	-0.01	-1.34
Specialisation			-0.12	-2.82*			-0.14	-2.11*
Household members					-0.01	-0.39	0.03	35.27*
Successor dummy					1.09	4.30*	0.17	41.77*
Education					1.17	2.80*	-0.02	-6.25*
Mills ratio					6.61	3.09*		

* indicates significance at the 10% level.

Before describing the estimation results we focus on the assumption that the group of farmers that supply off-farm labour is not a representative sample of all farmers. Based on this assumption we choose to use sample selection estimation approaches. The Mills ratio used for sample selection correction in the Wooldridge (1995) estimation approach is significant[5] at the 5% level. Therefore, we conclude that our sample selection assumption is correct. Now we know that a major assumption we made is justified, we can focus on the parameter estimates. From section 3 follows that the parameter estimates for off-farm wage; education level and age are biased in our estimated models. These variables control for the fact that we use a time invariant national off-farm wage rate instead of individual specific wages. The estimation

[5] When we say significant we of course mean significantly different from zero.

results for the other parameters are discussed below. We will discuss the estimation results of the two approaches separately.

7.1 Kyriazidou (1997)/FE

Based on our theoretical model, the milk price can only have an income effect on labour supply, since it only appears in the income equation (8). The parameter estimate for milk price in the off-farm equation supports an income effect. However, the sign of the corresponding parameter in the on-farm labour supply equation is unexpected. One might be inclined to conclude that milk output is not fixed on the short term as we assume in our theoretical model and that the found parameter estimates reflect a substitution effect of milk price. However, this implies that in case of a milk price increase all farmers would increase milk production. This is clearly impossible given the limiting national quota amount. The parameter estimates for other output price and cattle price have unexpected signs in the off-farm labour supply equation. Both an income and a substitution effect between on-farm and off-farm labour imply opposite effects than the ones found. The parameter estimates for other output price and cattle price are insignificant in the off-farm labour supply equation. Other income has an effect on on-farm labour supply but not on off-farm labour supply. This is unexpected, since we already found an income effect on off-farm labour supply through milk price. The parameters for quota are significant and have expected signs based on a substitution effect. More on-farm labour is supplied at the expense of off-farm labour. Land has a significantly positive effect on on-farm labour supply, but no effect on off-farm labour supply. Buildings have no effect on on-farm labour supply, but a significantly negative effect on off-farm labour supply. This is difficult to explain. Machinery does not have a significant effect on labour supply. The debt over asset ratio effect is insignificant for off-farm labour supply and significantly negative for on-farm labour supply. It is counterintuitive to work less if relative debt increases. The unemployment rate does not have a significant effect on labour supply. Specialisation in milk production has a significantly negative effect on on-farm labour input, as expected. Overall, many of the parameters estimated with Kyriazidou (1997)/FE have unexpected signs.

7.2 Wooldridge (1995)/Mundlak (1978)

The parameter estimates for milk price do not show strong evidence for an income effect. They are insignificant. However, the other income parameter is significantly negative in both equations. This implies that there is a negative income effect on labour. The fact that we do not find this through the milk price is caused by the limited variability in the milk price variable. This is both because we do not have farm specific milk prices and because the milk quota policy is partly introduced to reduce milk price variability. The parameter estimates for other output price are insignificant. This can be explained by the fact that all farms in the data set are specialised dairy farms and again little variability in the price variable. The effect of cattle price on off-farm labour supply is positive. This is in correspondence with a substitution effect between on-farm and off-farm labour supply. The complete substitution effect requires a negative effect of cattle price on on-farm labour supply. This effect is not found. However, cattle price also has a substitution effect between cattle and other input variables like on-farm labour, which is opposite to the substitution effect between on-farm and off-farm labour. From the insignificant value of the cattle price parameter in the on-farm labour supply equation, we conclude that neither of the two opposing effects is stronger. The parameters for quota are significant and have expected signs based on a substitution effect. Land has a significantly positive impact on on-farm labour supply. It also has a significantly positive effect on off-farm labour supply. This indicates that larger farms supply more off-farm labour and is in correspondence with other estimation results for off-farm labour supply (see e.g. Goodwin and Holt (2002) and Ahituv and Kimhi (2002)). Buildings have a significantly negative effect on off-farm labour supply. This is an unexpected result. An explanation is that redundant buildings are rented out and the found effect is an income effect. The maintenance of the buildings explains that this effect is not found for on-farm labour. Machinery does not have a significant impact on labour supply. The debt over asset ratio and the unemployment rate do not have a significant effect on labour supply. Specialisation in milk production has a significantly negative effect on on-farm labour input, as expected. Number of household members only has a significant effect on on-farm labour supply. Apparently, if there are household members they are easily expected and reported to work on-farm. Overall, most parameters estimated with Wooldridge (1995)/Mundlak (1978) have expected signs. Furthermore, it shows that time invariant variables are important for labour supply explanation.

7.3 Recapitulation

The two estimation methods result in different sets of parameter estimates. Based on a Hauseman test (see e.g. Greene, 2003) the statistical equivalence of the set of common parameters in the two off-farm labour supply equations is strongly rejected. We find a test value of 41.51, which clearly exceeds the 5% critical value for the χ^2_{12} distribution of 21.03. The equivalence of the set of common parameters in the two on-farm labour supply equations is hardly rejected. We find a test value of 22.45, which only just exceeds the 5% critical value for the χ^2_{13} distribution of 22.36. From an economic point of view the Kyriazidou (1997)/FE estimation approach gives implausible results, whereas this is not the case for the Wooldridge (1995)/Mundlak (1978) estimation approach. Kyriazidou (1997) is robust but less efficient and this may be the source of the discrepancy, but there is a need to explore further the differences in these two estimation approaches in these types of model.

8. Policy simulations

In this section we use the models estimated with the Kyriazidou (1997)/FE and the Wooldridge (1995)/Mundlak (1978) estimation approach to simulate the possible effects on labour supply of Dutch dairy farms of the CAP reform agreement of June 26, 2003. As base run we take the actual situation in 1999/00, the last year for which we have data. We calculate the effects as if the reform would be fully implemented in 1999/00, so we do not take the phased introduction or dynamic effects (e.g. structural changes) into account. Therefore, one could say that we do not pretend to give predictions but just provide information that is helpful to understand the effects of the 2003 CAP reform for Dutch dairy farming.

The three elements of the 2003 CAP reform are a milk price reduction, a quota increase and an introduction of direct income payments. Milk price and quota are explanatory variables in our estimated models. We simulate the effects of direct income payments by increasing the other income variable with the direct income payment.

We calculate the effects for the following scenarios:

1. S1a: CAP reform. For this simulation we assume a milk price reduction of 21%, this is based on the intervention price cuts in the CAP reform for skimmed milk powder and butter of 15% and 25% respectively. To determine the milk price we multiplied the intervention price reduction of skimmed milk powder with 0.4 and 0.6 for butter, as is done in the Mid Term Review proposals of the European Commission. The quota increase is 1.5%. Direct income payments equal 35.5 €/tonne.
2. S1b: As S1a but without income compensation.
3. S2a: see S1. Given the uncertainty about what the milk price will be after the CAP reform we assume a 15% price decrease. This can be considered as a minimum price decrease.
4. S2b: As S2a but without income compensation.

9. Simulation results

Below we describe the simulation results for both the Kyriazidou (1997)/FE and the Wooldridge (1995)/Mundlak (1978) estimation approach. Although we calculate for each individual farm in the sample the policy effects, we only present average changes. During the simulations we keep all other variables at their 1999/00 level.

9.1 Kyriazidou (1997)/FE

Simulation results show that labour supply is sensitive to the 2003 CAP reform. In S1a on-farm labour decreases 4.74% and off-farm labour increases 57.76% (see Table 2). Results are presented in percentage change. This explains why the change in off-farm labour is higher than the change in on-farm labour. However, the relation between on-farm and off-farm labour is not such that it can explain the big difference in change we find. Recall that these simulations are based on a model that did not meet some economic requirements. For this reason we do not trust these simulation outcomes.

Table 2: Simulation results based on Kyriazidou (1997)/FE

	S1a	S1b	S2a	S2b
On-farm labour supply	-4.74%	-4.36%	-3.38%	-3.08%
Off-farm labour supply	+57.76%	+57.68%	+38.25%	+38.18%

9.2 Wooldridge (1995)/Mundlak (1978)

Simulation results show that labour supply is rather insensitive to the 2003 CAP reform. In S1a on-farm labour increases 0.51% and off-farm labour decreases 1.05% (see Table 3). One has to take into account that the results are presented in percentage change and that off-farm labour supplying dairy farmers supply more on-farm then off-farm labour. Calculations suggest that on average on-farm labour supply is about 15% of total labour supply for off-farm labour supplying farms. This results in an increase of total labour supply in S1a. In S2a, where the milk price decrease is 15% instead of 21%, the income effect on labour supply is less. On-farm labour increases 0.29% and off-farm labour decreases 1.82%. Again, there is a shift from off-farm to on-farm labour supply. However, now total labour supply remains approximately the same. S1b and S2b show that both on-and off-farm labour supply increases if income is not compensated by a direct income payment.

Table 3: Simulation results based on Wooldridge(1995)/Mundlak(1978)

	S1a	S1b	S2a	S2b
On-farm labour supply	+0.51%	+0.95%	+0.29%	+0.73%
Off-farm labour supply	-1.05%	+0.49%	-1.82%	-0.30%

Direct income payments (farm payments) in the 2003 CAP reform are assumed decoupled from production decisions. The simulation results show an effect of the farm payments on on-farm labour supply. This effect is small and therefore the effect of farm payments on production through labour supply will be small. We conclude that based on these results we cannot reject the assumption that farm payments are decoupled from production decisions.

9.3 Recapitulation

The different estimation approaches lead to different estimates and therefore to different policy simulation results. The strong difference in simulation outcomes underlines the need to explore further the differences in the two estimation approaches.

10. Summary and conclusions

This research focuses on the estimation of on- and off-farm labour supply equations for Dutch dairy farmers that are suitable for 2003 CAP reform simulations. Off-farm labour supply of Dutch dairy farmers is characterised by the fact that only half of the farm households supply off-farm labour. We assume that the farmers that supply off-farm labour are not a representative sample from the Dutch dairy farmers' population. This results in a sample selection estimation problem that has to be taken into account in estimation. To this end we use the panel data sample selection estimation approaches of Kyriazidou (1997) and Wooldridge (1995) to estimate the off-farm labour supply equation. The latter approach is based on the fixed effect panel data estimation approach for linear models of Mundlak (1978). The FE and Mundlak approaches are used to estimate the linear on-farm labour supply equation. The two estimation methods result in different sets of parameter estimates. From an economic point of view the Kyriazidou (1997)/FE estimation approach gives implausible results, whereas this is not the case for the Wooldridge (1995)/Mundlak (1978) estimation approach. Kyriazidou (1997) is robust but less efficient. This may be the source of the discrepancy, but there is a need to explore further the differences in these two estimation approaches in these types of model. Estimation results show an income effect and a substitution effect between on-farm and off-farm labour. The different estimation approaches lead to different policy simulation results. The strong difference in simulation outcomes underlines the need to explore further the differences in the two estimation approaches. Based on Wooldridge (1995)/Mundlak (1978) we find that the effects of the rather large change due to the 2003 CAP reform are relatively small for labour supply of Dutch dairy farmers.

The results of our study are obviously subject to some qualifications. The model used for simulation can be characterised as a comparative static short-term model, since technology, most production factors (capital, land and labour) and prices of variable inputs are assumed fixed and no explicit time path for the changes is given. In the longer term factors and variable input prices are no longer fixed and alternative technologies may come available. Moreover, it is unclear what the effect of 2003 CAP reform on the milk price will be, estimations in the Netherlands vary between 15% and 21%. We do not take into account farm continuation problems that might arise given the large decrease in profits. Making the model dynamic and including environmental policies could be interesting topics for future research. The model presented here can serve as a building block in this type of extended analysis.

References

Ahituv, A. and A. Kimhi (2002). Off-Farm Work and Capital Accumulation Decisions of Farmers Over the Life-Cycle: The Role of Heterogeneity and State Dependence. *Journal of Development Economics* 68 329-353.

Baltagi, B.H. (2001). *Econometric Analysis of Panel Data*. Chicester: John Wiley and Sons, Ltd.

Bierens, H.J. (1987). Kernel estimators of Regression Functions. In *Advanced Econometrics: Fifth World Congress*, Vol. 1, ed. By T.F. Bewley. Cambridge: Cambridge University Press.

European Commission (2003). Press Release: EU fundamentally reforms its farm policy to accomplish sustainable farming in Europe. Brussels: European Commission DN:IP/03/898 Date: 26/06/2003. (http://europa.eu.int/comm/agriculture/).

CBS/LEI-DLO (2000). *Landbouwcijfers*. Landbouw-Economisch Instituut/Centraal Bureau voor de Statistiek, 's-Gravenhage/Voorburg.

Chamberlain, G. (1982). Multivariate Regression Models for Panel Data. *Journal of Econometrics* 18, pp. 5-46.

Goodwin, B.K. and M.T. Holt (2002). Parametric and Semiparametric Modeling of the Off-Farm Labor Supply of Agrarian Households in Transition Bulgaria. *American Journal of Agricultural Economics* 84 pp. 184-209.

Greene, W.H. (2003). *Econometric Analysis*. London: Prentice-Hall international.

Heckman, J.J. (1979). Sample Selection Bias as a Specification Error. *Econometrica* 47 pp. 153-161.

Huffman, W.E. (1980). Farm and Off-Farm Work Decisions: The Role of Human Capital. *Review of Economics and Statistics* 62, pp. 14-23.

Huffman, W.E. and H. El-Osta (1998). Off-Farm Work Participation, Off-Farm Labor Supply and On-Farm Labor Demand of U.S. Farm Operators. *Selected Paper for the 1998 Summer AAEA Meeting*: Salt Lake City, UT.

Kyriazidou, E. (1997). Estimation of a Panel Data Sample Selection Model. *Econometrica* 65, pp. 1335-1364.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica* 46 pp. 69-85.

Weiss, C.R. and W. Briglauer (2000). Determinants and Dynamics of Farm Diversification. *FE-Working Papers*. Kiel: University of Kiel.

Woldehanna, T., A. Oude Lansink and J. Peerlings (2000). Off-Farm Work Decisions on Dutch Cash Crop Farms and the 1992 and Agenda 2000 CAP Reforms. *Agricultural Economics* 22 pp. 163-171.

Wooldridge, J.M. (1995). Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions. *Journal of Econometrics* 68 pp. 115-132.

Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*. Cambridge and London: MIT Press.

Appendix A: Supply quota

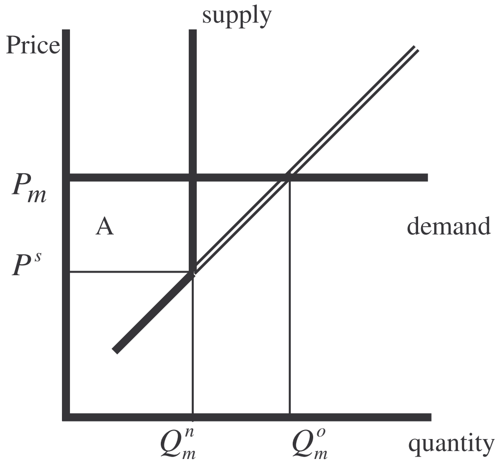

Figure A.1 Supply quota

Quota rent: A

In the case of a supply quota Q_m^n and market price P_m, the shadow price of production P^s gives the marginal costs of production. The market price is the reward for the production right (quota) and the factor inputs supplied by the farm household (labour and capital). The shadow price of the quota equals $P_m - P^s$ and is the reward for the production right. The shadow price of production is the reward for the factor inputs labour and capital. Figure A.1 shows that the shadow price of production does not change with a change of the output price. From this follows that the marginal products of the factor inputs are not dependent on the market (output) price.

Appendix B: Data

Table B1: Data for average specialised dairy farm in the Netherlands in 1999/00.

Variable	Dimension	Mean	Standard deviation
Off-farm income (> 0 for 46.2% of observations)	1000 Euro	5.087	7.765
On-farm labour	Hours	4068	1484
Off-farm wage index	1991 = 100	105.46	10.74
Milk price index	1991 = 100	98.44	5.15
Other output price index	1991 = 100	100.08	11.11
Input price index	1991 = 100	100.42	4.71
Other income	1000 Euro	10.291	8.821
Quota	1000 Kilo	444.547	277.351
Land	Hectares	35.264	19.820
Buildings	1000 Euro	171.216	105.044
Machinery	1000 Euro	77.371	51.463
Debt	Percentage of Assets	27.45	18.02
Unemployment rate	Percentage	6.19	1.28
Specialisation in milk	Percentage	75.39	9.43
Household members	Number	4.65	1.91
Successor	Percentage	41.67	
Education	Dummy	2.45	0.60

Appendix C: Off-farm labour decision estimation results

Table C1: Estimation results for off-farm labour supply decision.

Estimation approach:	Fixed effects logit		Pooled probit	
	Estimate	t-statistic	Estimate	t-statistic
Constant			3.67	0.98
Log off-farm wage	0.40	0.80	0.18	0.48
Milk price	2.72	2.42*	0.97	1.19
Other output price	-0.68	-1.22	-0.25	-0.64
Cattle price	2.88	2.66*	1.19	1.54
Other income	-1.70	-0.72	-0.89	-0.53
Quota	-0.43	-0.87	-0.47	-1.11
Land	0.87	1.36	0.45	0.85
Buildings	-0.63	-1.63	-0.29	-1.01
Machinery	-0.46	-0.50	-0.18	-0.26
Debt over asset ratio	0.45	1.08	0.27	0.90
Unemployment rate	0.04	1.20	0.01	0.57
Specialisation	-0.07	-0.09	0.01	0.02
Number of household members			0.01	1.11
Successor dummy			0.13	3.71*
Education			0.31	10.96*

* indicates significance at the 10% or smaller level.

Table C.1 gives the parameter estimates for the off-farm labour decision equation for both the fixed effects logit estimator used in the Kyriazidou (1997) estimation approach and the pooled probit estimator used in the Wooldridge (1995) estimation. These results are not of primary interest in this paper. Therefore, we do not describe them extensively. The overall impression is that there are not many significant parameters. With the fixed effects logit estimator, only milk and cattle price are significant. With the pooled probit estimator only the presence of a successor and education are significant. None of the parameters associated with time varying variables are significantly different from zero.